CURTIS, MALLET-PREVOST, COLT & MOSLE LLP

ALMATY
ASHGABAT
ASTANA
BEIJING
BUENOS AIRES
DUBAI
FRANKFURT
HOUSTON
ISTANBUL
LONDON
MEXICO CITY
MILAN
MUSCAT
NEW YORK
PARIS

ATTORNEYS AND COUNSELLORS AT LAW

1717 PENNSYLVANIA AVENUE, N.W.
WASHINGTON, D.C. 20006

TELEPHONE 202-452-7373
FACSIMILE 202-452-7333
WWW.CURTIS.COM

WRITER'S DIRECT:
TEL.: 202-452-7340
E-MAIL: DPORTER@CURTIS.COM

June 23, 2014

PUBLIC DOCUMENT

The Honorable Penny Pritzker
Secretary of Commerce
U.S. Department of Commerce
Attn: Import Administration
APO/Dockets Unit, Room 1870
14th Street & Constitution Avenue, N.W.
Washington, D.C. 20230

Re: LG Electronics' Comments on Differential Pricing Methodology

Dear Madam Secretary:

With this submission LG Electronics, Inc., and LG Electronics USA, Inc., (collectively "LGE") hereby provide their comments in response to the Commerce Department's request for comments regarding the continued development of a methodology for determining whether the legal criteria set forth in Section 777A(d)(1)(B), 19 U.S.C. 1677f-1(d)(1)(B), of the Tariff Act of 1930, as amended, have been satisfied and thereby justify use of an alternative comparison methodology for calculating antidumping ("AD") margins. *See* Differential Pricing Analysis, Request for Comments, 79 Fed. Reg. 26,720 (May 8, 2014).

LGE appreciates the opportunity to provide these comments to the Department. The comments are enclosed.

Respectfully submitted,

/s/ Daniel L. Porter

William H. Barringer
Daniel L. Porter
James P. Durling
Ross Bildingmaier
Anya Naschak, Senior Trade Advisor
Paul Casas, Trade Advisor

Before the United States Department of Commerce
International Trade Administration

COMMENTS OF LG ELECTRONICS INC. AND LG ELECTRONICS USA INC.

Concerning Suggested Changes to the Commerce Department's Differential Pricing Methodology

William H. Barringer
Daniel L. Porter
James P. Durling
Ross Bidlingmaier
Anya Naschak, Senior Trade Advisor
Paul Casas, Trade Advisor

Curtis, Mallet-Prevost, Colt & Mosle LLP
1717 Pennsylvania Ave., N.W.
Washington, D.C.

June 23, 2014

Table of Contents

Page

INTRODUCTION

These comments are submitted on behalf of LG Electronics Inc. and LG Electronics USA Inc. (hereinafter "LGE") in response to the Commerce Department's request for comments regarding the continued development of a methodology for determining whether the legal criteria set forth in Section 777A(d)(1)(B), 19 U.S.C. 1677f-1(d)(1)(B), of the Tariff Act of 1930, as amended, have been satisfied and thereby justify use of an alternative comparison methodology for calculating antidumping ("AD") margins. *See* Differential Pricing Analysis, Request for Comments, 79 Fed. Reg. 26,720 (May 8, 2014). LGE appreciates the opportunity to provide these comments to the Department.

These comments consist of five main sections as summarized below:

- The first section explains why the Department should return to its prior targeted dumping regulation, which more faithfully reflects the statutory language and sound policy choices than does the current differential pricing policy.
- The second section urges the Department to adopt a proper definition for the statutory term "significantly", in particular by recognizing that "significantly" has both quantitative and qualitative aspects and by allowing respondents to provide explanations as to why prices differ.
- The third section urges the Department to abandon utilization of the "zeroing practice" when applying the alternative comparison methodology after finding the criteria of 19 U.S.C. 1677f-1 have been met.
- The fourth section urges the Department to explain more clearly why averaging cannot take into account price differences.
- And the fifth section concludes the comments by providing a number of specific and somewhat more technical comments about the new differential pricing policy, and how it could more faithfully reflect the statutory language and sound policy choices.

I. THE DEPARTMENT SHOULD RECONSIDER ITS DECISION TO WITHDRAW ITS TARGETED DUMPING REGULATION

The Department should take this opportunity to reconsider its decision, published in the Federal Register on April 22, 2014, to no longer apply the regulations that previously governed targeted dumping.[1] Recognizing the Court of International Trade's decision in *Gold East (Jiangsu) Paper Co. v. United States*, which found withdrawal of the targeted dumping to be procedurally defective under the Administrative Procedures Act ("APA"), the Department undertook full notice and comment to determine "whether to reinstate the regulations or to continue to treat them as withdrawn."[2] In the published final notice, the Department determined "to continue not to apply the withdrawn targeted dumping regulations in less-than-fair-value investigations."[3]

LGE submits that the *current* request for comments provides the Department with an opportunity to remedy the *substantive* defect in the Department's decision to no longer apply the withdrawn regulation. Leaving aside for present purposes the Department's procedural compliance with the APA in not applying an improperly withdrawn regulation, the decision to not apply the targeted dumping provision does not comply with the APA's requirement that an agency's action be rational in light of the available evidence. Furthermore, even if the Department is not legally required to apply this APA requirement, the interests of good policy demand application of the principles set forth in the prior targeted dumping regulation.

---

1 77 Fed. Reg. 22,371 (April 22, 2014) ("Final Notice of Non-Application").
2 *Id.*
3 *Id.*

A. The Department's Failure to Satisfy the Requirements of the Administrative Procedure Act Requires Continued Application of the Targeted Dumping Regulation

In addition to procedural obligations, the APA imposes other constraints on the regulatory actions that can be taken by an agency. Although an agency has clear discretion to change a policy when circumstances require, the change must be rational in light of the evidence before the agency. This requirement has particular force where an agency is seeking to rescind a properly promulgated regulation – done in full compliance with the notice and comment requirements of the APA – that benefited from active participation and comments from interested parties. The Supreme Court explained that "an agency changing its course by rescinding a rule is obligated to supply a reasoned analysis for the change beyond that which may be required when an agency does not act in the first instance." The agency must provide a reasoned analysis, based on evidence, to justify the withdrawal (or non-application) of a properly promulgated regulation.

The Department has failed to provide the required explanation for why non-application of the targeted dumping regulation is necessary or appropriate. The Department has cited the need for flexibility to make targeted dumping determinations on a case-by-case basis as the rationale for non-application of the regulation, as the prior regulations may have made it more difficult to unmask targeted dumping.[4] There are two problems with this reasoning.

First, the Department has provided no explanation for how the targeted dumping regulation prevented the Department from making determinations on a case-by-case

[4] Final Notice of Non-Application, 77 Fed. Reg. at 22,374-75.

basis. The regulations included only general principles on how targeted dumping should be administered, and did not in any way limit the Department's ability to choose one particular methodology over another. The very discretion that the Department claims it needs in fact already existed under the prior targeted dumping regulation. Non-application of the regulation does not in any way advance this stated purpose.

Second, the Department has provided no evidence that the targeted dumping regulations masked or could mask targeted dumping. Given that the Department effectively found the existence of targeted dumping in numerous cases under the withdrawn regulation, it bears the burden of demonstrating that the regulation somehow prevented the Department from unmasking targeted dumping in specific cases. The Department has not explained how a regulation that sets forth two general principles, including the requirement that any methodology used employ standard statistical techniques, could mask targeted dumping. The Department has not explained how non-application of the targeted dumping serves the purposes set forth in the published notices.

The Department thus has not satisfied the requirements of the APA. An agency's action must be supported by evidence and a reasoned explanation. Because the Department is not able to do so, LGE urges the Department to reconsider its decision to not apply the targeted dumping regulation.

B. Sound Policy Reasons Require the Department to Continue to Apply the Targeted Dumping Regulation

LGE submits that regardless of the APA requirements discussed above, the Department should continue to apply its prior targeted dumping regulation as a matter of good policy. As explained above, the regulation in no way restricted the Department's ability to adopt a new methodology to address targeted dumping. Rather, the regulation simply set forth broad, commonsense principles within which the Department would be free to make case-by-case decisions that reflect the merits of a particular case. Changes to the methodology on a case-by-case basis could continue, but within the broad framework set by the regulation. LGE believes that the regulation strikes the proper balance between providing some level of predictability to interested parties and providing the Department with the discretion to continue to adapt its targeted dumping methodology.

In adopting the targeted dumping regulation in 1997, the Department cited the need for "predictability and transparency" in promulgating the requirement that the Department use statistically valid techniques to determine the existence of targeted dumping.[5] The need for such predictability and transparency continues to this day. Notably, the regulation did not impede on the Department's ability to implement several different tests before withdrawal of the regulation. The Department adopted a series of methodologies on a case-by-case basis, which is precisely the discretion that it now purportedly seeks through withdrawal of the regulation. As demonstrated by the multiple

[5] 62 Fed. Reg. at 27347 (Dep't of Commerce, May 19, 1997).

prior tests (including the *Pasta* and *Nails* tests), such discretion exists under the prior regulation. The Department can continue to exercise its discretion within the parameters of the regulation.

The regulation's requirement that the alternative methodology of using the average-to-transaction price comparison for only sales found to be part of the targeted dumping likewise represents a sound policy decision that should be reinstated. The Department adopted this position after extensive comments from interested parties. This process led the Department to conclude that to apply the alternative comparison methodology to all transactions would "be unreasonable and unduly punitive."[6] This finding reflects the common sense conclusion that application of any targeted dumping remedy should be relative to the extent of the targeted dumping. The provision still provides the Department with the discretion to apply an alternative comparison methodology to additional sales, but merely sets the starting point for the analysis. The Department itself concluded that any other starting point would be unreasonable. The provision struck a proper balance between providing predictability on the remedy and granting the Department the discretion to deviate from the "normal" remedy.

The targeted dumping regulation did not set forth detailed requirements on how the Department is to perform its targeted dumping analysis. Rather, it balanced the interests of predictability with the Department's discretion to develop a methodology through broad, general principles to be followed. There is no sound policy reason to discard these principles, which were the product of extensive comments by interested

[6] 62 Fed. Reg. 27375 (Dep't of Commerce May 19, 1997).

parties and reflect common sense guidelines for a targeted dumping analysis. LGE believes that the Department should take this opportunity to clarify that it will continue to apply the targeted dumping regulation.

II. THE DEPARTMENT SHOULD ADOPT A PROPER INTERPRETATION OF THE STATUTORY TERM "SIGNFICANTLY"

As made clear by the statutory language itself, 19 U.S.C. § 1677f-1.2 requires not only that the prices examined differ, but also that they differ "significantly." The word "significant" or "significantly" means more than just "large." Indeed, the word "significant" conveys both qualitative and quantitative aspects. LGE urges the Department to reflect this understanding in implementing the statutory language. And LGE further submits that adopting a proper interpretation of "significantly" requires allowing respondents to provide explanations as to *why* prices differ. We address each below.

A. Significant Means More Than Just Large

The first definition of the word "significantly" provided by the New Shorter Oxford English Dictionary is "having or conveying meaning." While the English word may also have quantitative aspects,[7] there is no doubt that the qualitative aspects of the word are at least as important, if not more so, than the quantitative aspect.

The ordinary meaning of the word "significant" as encompassing more than merely quantitative consideration is also evident by the applicable statutory context and

[7] Merriam Webster contains a definition that includes "large enough to be noticed or have an effect." The same dictionary, however, includes the definition "having a special or hidden meaning."

purpose of the relevant provision. With respect to the statutory context and purpose of the phrase "differs significantly," the relevant sentence is an exception to the general rule that the authority compare average prices in both the export and normal value markets (or individual transaction prices in both markets). When average prices are used, quantitative differences among individual export prices, even large differences, are eliminated by the averaging process. When individual prices are examined, on the other hand, it would not be unusual to find large variations in those prices. Large individual price differences in this context would thus likely be perfectly normal. The fact that the text uses the word "significantly," rather than "large" to describe the differences in prices in this context supports the conclusion that the drafters of the provision intended the word to mean something more than merely large differences. And therefore, the use of the word "significantly" in the context of 19 U.S.C. § 1677f-1 suggests that it was intended to refer to differences that are "meaningful," not merely large, emphasizing the qualitative aspects of the term.

The underlying purpose of the AD law also confirms the importance of the qualitative aspect of the word "significantly" in 19 U.S.C. § 1677f-1. The AD law requires that there be a "fair comparison" between export price and normal value. Again, since the normal rule set forth in 1677f-1 requires use of both average export price and average normal value in most cases, quantitative differences in individual prices will balance out in the averaging process and a fair comparison can be made in most cases. However, where individual export prices are compared to average normal value under the exception (set forth in 19 U.SC. § 1677f-1), otherwise explainable differences in

individual prices could result in comparisons that are not "fair," since numerical differences would be smoothed out in the calculation of normal value but not in the export price.

In this context, it is apparent that the use of the word "significantly" to describe price differences must mean something other than merely large quantitative differences. It must mean price differences that are meaningful enough in some particular situations that averaging them in with other export prices would not lead to a "fair comparison."

Despite the obvious broader meaning of the word "significant" that requires consideration of the qualitative aspects of price differences, the Department has repeatedly invoked the exception set forth in 19 U.S.C. § 1677f-1 based on exclusively numerical criteria, rather than holistically evaluating both quantitative and qualitative dimensions. Indeed, the Department's differential pricing test sets forth purely quantitative criteria. The Department then applies these tests mechanically, and then analyzes only the quantitative differences among those average prices, never examining the reasons for any price differences. To the contrary, when the respondents suggest other reasons for the differences in export prices, the Department typically refuses to consider those reasons.

LGE submits that the use of exclusively numerical tests to determine whether price differences are "significant" violates the ordinary meaning of the term.

B. Analyzing "Significantly" Requires Allowing Respondents To Provide Reasons Why Prices Differ

Based on numerous recent final determinations, it is clear that the Department applies its current differential pricing test mechanically and does not allow any consideration of *why* prices vary. Such mechanical application is not only wrong as policy matter, it is also contrary to the AD statute.

When the AD statute uses the terms "significantly" and "pattern" it does mean just that a difference is large and a certain number of transactions. A significant difference has to be large, but it also has to be meaningful in some specific context. A pattern is not just a number of transactions, but a number of transactions that show something. Thus, in the context of targeted dumping, significant differences and the pattern must be meaningful for the effort to identify possible targeted dumping. Differences that occur for some natural business reasons are thus not "significant" in this context. Such differences have been explained and they are not targeted dumping. Not taking into account the reasons for price differences means the Department's test is flagging as targeted dumping many instances that cannot possibly be targeted dumping.

Consider two real world examples. First, consider the simple example of prices that are changing over time because costs are falling or rising. With changing costs it is essentially impossible for prices not to change, at least somewhat. Yet these price changes then trigger a finding of differential pricing. Prices to a single customer may change over time. Prices to different customers may vary depending on how much each customer bought at different points in time. Prices to different regions may vary

depending on which customers bought during what point in time. In none of these situations does the price variation reflect anything other than normal business practices.

Or consider the example of foreign exporters matching the business practices of its U.S. competitors. If the U.S. companies have a practice of steep discounts during a holiday period, foreign companies have little choice but to follow this business practice set by the domestic firms. Again this situation does not reflect any targeting, but rather the reaction to pricing set by the domestic industry itself.

The facts of each case will be different. In some cases the Department might find the explanation reasonable and persuasive; in other cases the Department might find the explanation insufficient. But the point is that this determination should be made in each case based on the explanations offered and the degree to which the explanation can be confirmed with record evidence in that case. It makes little sense to adopt a fixed rule at the outset that the reasons for a price difference never matter.

And so, LGE submits that Commerce should allow parties to rebut any preliminary finding of differential pricing by showing a neutral business justification (such as falling costs) that explains the price variance. Price variance for a normal business reason unrelated to any possible "targeted dumping" should not meet the definition of a proper test for differential pricing.

III. THE DEPARTMENT SHOULD RECONSIDER ITS DECISION TO APPLY ZEROING WHERE IT FINDS THE EXISTENCE OF DIFFERENTIAL PRICING

Under its current approach, whenever the Department finds that more than the sales prices for more than 33 percent of U.S. sales transactions meet the differential

pricing criteria, the Department employs a remedy that utilizes different forms of its zeroing practice. LGE submits that the Department should abandon the application of the long-discredited zeroing practice.

A. Commerce Should Adhere to Clear U.S. International Obligations and Eliminate the Use of Zeroing When Applying the Differential Pricing Remedy

Commerce currently applies the zeroing methodology to at least a portion of all sales where (1) the statutory criteria for targeted dumping are satisfied and (2) the ratio of sales that pass the Cohen's *d* test exceeds 33 percent. Specifically, where the statutory criteria are satisfied and Commerce determines that the average-to-average comparison methodology cannot account for price differences, it applies the average-to-transaction methodology (and zeroing) to only those sales that pass the test where that percentage is between 33 and 66 percent, and applies the average-to-transaction methodology to all sales where that percentage is greater than 66 percent. Commerce's decision to apply zeroing to these sales violates a long line of WTO precedent. It serves all parties' long-term interests for Commerce to adopt a policy that complies with U.S. international obligations.

The WTO jurisprudence on the use of zeroing in antidumping proceedings is well-established. Article VI:1 of the General Agreement on Tariffs and Trade 1994 and Article 2.1 of the Anti-Dumping Agreement define the concepts of "dumping" and "margin of dumping" under the covered agreements. The text of both make clear that the determination of dumping must be made with respect to the product as a whole, and not

any subset of the product under investigation. While this definition does not prevent an authority from undertaking multiple comparisons using groups or models, the results of the multiple comparisons at the sub-product level are not "margins of dumping." Rather, those results reflect only intermediate calculations made by an authority in the context of establishing the margin of dumping for the product under investigation. Only by aggregating the intermediate values can an authority establish a margin of dumping for the product as a whole.

Beginning with the decision in *EC – Bed Linen*, the Appellate Body has repeatedly and consistently held that dumping margins must be calculated with respect to the "product as a whole." As explained in *US – Softwood Lumber V*, "dumping , within the meaning of the Anti-Dumping Agreement, can therefore be found to exist only for the product under investigation as a whole, and cannot be found to exist only for a type, model, or category of that product." An authority may not selectively use some prices when determining the dumping margin while excluding other prices.

Under Commerce's current practice, a subset of all sales – those that produce a negative dumping margin – are systematically eliminated from the final dumping margin calculation. Where Commerce applies the average-to-transaction comparison methodology and the percentage of sales that pass the ratio test exceeds 33 percent, certain transactions are excluded from the calculation. Based on clear WTO precedent, this current practice is in violation of the United States' obligations. Commerce should bring its practice into conformity with those obligations and eliminate the use of zeroing.

B. If Commerce Utilizes The Zeroing Methodology, It Should Only Do So For Sales Found To Be Differentially Priced

As explained above, Commerce's current practice applies the zeroing methodology to all sales where it determines that more than 66 percent of U.S. sales have met the criteria. Under this scenario, Commerce applies zeroing even to those sales that do not meet the statutory criteria of targeted or differentially priced sales. Thus, some sales, even though they were not found to be differentially priced, are arbitrarily eliminated from the dumping margin calculation. Commerce should revise this practice, as it is contrary to both the U.S. antidumping statute and the United States' international obligations.

The antidumping statute specifies the methodology for calculating dumping margins and requires Commerce to apply the exceptional comparison methodology only to those sales found to qualify for the exception. Commerce's current practice ignores several critical elements of the statute. First, Commerce has allowed the discretionary to trump the mandatory. The statute provides a limited exception to using the standard average-to-average methodology. That exception can only be triggered when the conditions set forth in the statute have been satisfied. Commerce has expanded the scope of this exception beyond the language in the statute through applying the exception to sales that have not satisfied the necessary conditions.

Second, Commerce's current practice ignores the statutory requirement that Commerce explain why "such differences" in prices cannot be taken into account using the standard methodology. There is no rationale for applying an alternative methodology

to transactions without "such differences" in pricing. The alternative methodology is intended to address an issue that is specifically missing in the non-differentially priced sales.

LGE's suggested change is consistent with Commerce's prior stance on this very issue. In promulgating the targeted dumping regulation, Commerce cited the statutory language as a reason for it to "apply the average-to-transaction approach solely to address the practice of targeted dumping," and apply the exception "exclusively to those sales in which the criteria for determining targeted dumping are satisfied."[8] The text of the statute that formed the basis of Commerce's past conclusion has not changed. Nor has the practical consideration that applying the alternative methodology to non-differentially priced sales can "in many instances {} be unreasonable and unduly punitive," another finding previously made by Commerce.[9] The underlying facts have not changed, and nor should have Commerce's policy.

In addition to the U.S. legal and policy reasons supporting LGE's suggested change, the United States' international obligations likewise compel Commerce to do so. Article 2.4.2 of the Antidumping Agreement, like the antidumping statute, presumes the use of the average-to-average comparison methodology. An authority can deviate from this methodology only under specifically defined circumstances, as found in the second sentence of Article 2.4.2. But the Antidumping Agreement makes clear that the exception does not apply to those transactions not found to meet the criteria. Nothing in

[8] 62 Fed. Reg. 27,296, 27,375 (May 19, 1997).
[9] 61 Fed. Reg. 7,308, 7,350 (Feb. 27, 1996).

the text of the Antidumping Agreement or any other agreement permits Commerce to apply the exceptional average-to-transaction methodology to export transactions that do not meet the criteria. Commerce's current practice directly contravenes the clear text of the Antidumping Agreement.

C. The Department Should Also Abandon Its Practice of Not Allowing Negative AD Margins for Sales Transactions Not Found To Satisfy the Differential Pricing Criteria To Offset Those That Do.

In addition to employing the two actual zeroing practices noted above, under its current approach, the Department also employs what is essentially a zeroing practice when it calculates the final AD rate for an investigated foreign exporter. Specifically, whenever the Department determines more than 33 percent, but few than 66 percent, of the U.S. sales transactions have sales prices that meet the differential pricing criteria, the SAS code utilized by the Department actually calculates two separate AD margins: (1) an AD margin for those transactions for which the Department has found satisfied the differential pricing criteria (for which zeroing is employed) and (2) and AD margin for those transactions for which the Department did not meet the differential pricing criteria (for which no zeroing is employed.) The SAS code then combines the two AD margins in order to generate a single AD rate for the investigated exporter. However, when it does so, if the second AD margin (for transactions not meeting the differential pricing criteria) is negative, such negative AD margin is not allowed to offset whatever positive AD margin exists for the first AD margin (for transactions satisfying the differential pricing test.)

LGE submits that such practice effectively constitutes zeroing that the WTO Appellate Body has condemned. And so, LGE urges the Department to abandon this practice.

IV. THE DEPARTMENT SHOULD EXPLAIN MORE CLEARLY WHY AVERAGING CANNOT TAKE INTO ACCOUNT PRICE DIFFERENCES

The requirement that Commerce provide an "explanation" in 19 U.S.C. §1677f-1(d)(i)(B)(ii) is not a requirement that Commerce provide any old general explanation of its actions. Rather, it is an explicit requirement that Commerce explain specifically why the "significant" price differences in the targeted exports "cannot be taken into account" by the average-to-average price comparison methodology. The Department has responded to this argument by saying that it provides such an explanation by simply stating that there were differences in the margins using the average-to-average methodology as compared to the average-to-transaction methodology. However, this perfunctory statement does not constitute a reasonable explanation.

First, the mere fact that Commerce found overall dumping margins to be different under the two methodologies says nothing about the export transactions that fell within the "prices that differ significantly…" as required by the statute. An explanation that the overall margin was different using the average-to-transaction methodology fails utterly to explain why the differences among the targeted transactions "could not be taken into account" using the average-to-average methodology, as required by the statute.

Perhaps more significantly, the mere fact that Commerce found a difference does not in any way prove that or explain why the "significantly" different pricing of exports "cannot be taken into account." The Department ignores the fact that it applied its remedy for differential pricing– so-called "zeroing" or setting all negative dumping margins equal to zero – when it applies the average-to-transaction methodology. When the Department applies this particular remedy under the average-to-transaction methodology, the resulting *overall margins will always, necessarily, be higher than they are under the average-to-average methodology*, so long as there is even one export sale priced above normal value. Consequently, a different result inevitably occurs from the remedy, because the "negative margins" on any above-normal value sales are set to zero and are not allowed to offset the positive margins on any below-normal value sales.

Yet the Department is not required by statute to use this particular remedy on any sales to which it applies its average-to-transaction methodology. Essentially, the Department is justifying its use of this specific remedy, by applying the remedy and then comparing the outcomes. But this logic is circular. The Department does not explain why the average-to-average transaction methodology "cannot take into account" the price differences it found to be significant when it merely announces that the margins are higher using that methodology, because it fails to explain to what extent the higher margins were purely the result of setting negative margins equal to zero, rather than the result of the price differences themselves.

As importantly, based on past cases, the Department provides no explanation as to how or in what way the differences in margins under the two methodologies were

"meaningful." Commerce does not, for example, state that the margins were more than X percent, and therefore were significant; nor does Commerce mention any other circumstances (such as the nature of price competition in the market) to shed light on why the differences it found would be meaningful. Thus Commerce's explanation of the difference is little more than an *ipse dixit* statement: the differences in margins are meaningful "because I said so."

It must be remembered that the statute does not require Commerce to explain why the differences in margins under the two methodologies are meaningful. It requires Commerce to explain why the average-to-average methodology *"cannot take" the "significant differences" in prices into account.* A statement that there are "meaningful" differences between the results of the two methodologies does not even begin to fulfill the requirement of the statute. As noted above, for example, if Commerce applied the average-to-transaction methodology without zeroing, there might well be no difference in the margins under the two methodologies. Thus, use of an average-to-average methodology might well take the significant differences in export prices fully into account. Commerce's "explanation" is thus not an explanation at all. It is wholly unresponsive to the requirement set forth in the statute.

V. THE DEPARTMENT SHOULD IMPLEMENT IMPORTANT CHANGES TO ITS DIFFERENTIAL PRICING METHODOLOGY

If the Department declines to return to its prior targeted dumping regulation, there are a number of specific changes that should be made to the new differential pricing

policy so as to better reflect the statutory language and sound policy choices. In particular, the Department should:

- Apply the Cohen's d test in conjunction with a recognized test of statistical significance, so as to ensure the measured "large" is in fact a meaningful rather than a spurious conclusion.
- Raise the threshold from 0.8 standard deviations to a higher threshold, preferably 1.96, so as to ensure the prices truly are different from the average and range of normal prices.
- Limit the differential pricing test to lower priced transactions, so as to ensure the test focuses on possibly targeted transactions.
- Test the three bases separately, and not aggregate the results into a single number, so as to ensure the 33% and 66% threshold are meaningful tests unbiased by the other basis.
- Stop excluding the test transactions from the base used for comparison, so as to ensure the comparison group is truly an accurate measure of the normal transactions.
- Switch to a weighted average variance, where applicable, so that different sample sizes and population structures return actual, meaningful results.
- Adjust the 25% test for when averaging can capture differences, clarifying that changes in the margin that are *de minimis* can be taken into account with averaging, so as to recognize that 25% of a small number is still a small number than should be ignored as *de minimis*.
- Require a larger minimum number of transactions, preferably at least 30 transactions, so as to ensure the means and standard deviations used in the differential pricing test are meaningful and not distorted or biased by outliers.
- Articulate clear definitions for the terms "region," "time period," and "customer" to provide respondents with some degree of predictability.

We discuss each of these specific suggestions in more detail in the sections below. These suggestions share the common goal of modifying the differential pricing test to ensure

that it identifies only prices that form a pattern of prices that differ significantly in some meaningful way.

Suggested Change #1: **Apply the Cohen's d test in conjunction with a recognized test for significance**

As noted in its *Federal Register* notice seeking comments, the Department is now analyzing differential pricing by using a new measure, called the "Cohen's d test." The Department characterizes this new "Cohen's d test" as a "generally recognized statistical measure." Yet this characterization is incorrect, and the use of the Cohen's d test takes the Department even further afield from a statistically sound approach to identifying any possible targeted dumping. The Department is trying to create the illusion of statistically validity, when in fact the Department is ignoring the entire issue of statistical significance.

The Cohen's d test is not an accepted measure of statistical significance, and no commentator would consider it a measure of statistical significance. Rather than test statistical significance – in other words, testing whether the conclusion being drawn is likely, or simply reflects the variance or random "noise" in the data – the Cohen's d test simply measures and standardizes the size of a difference between two mean values. That is what the Department means when it explains the Cohen's d test is a "measure of the extent of the difference between the mean of a test group and the mean of a comparison group." *Id.* at page 3. It measures "the extent of the difference," and is a common measure of the so-called "effect size." Cohen's d test is nothing more,

notwithstanding the Department's attempt to cloak this measure with more meaning than it has. Cohen's d test is not a measurement of statistical significance.

One purpose of the Cohen's d test is to measure the size of the difference. In other words, this procedure tries to address whether the difference between two mean values is "small," "medium," or "large." Such comparisons depend inherently on having some convention on "small" or "large" relative to what. The Cohen's d test simply adopts the convention that "small" or "large" can be measured relative to the standard deviation of the population being studied (or an estimate of that standard deviation for the data being studied). Thus, to say that the Cohen's d test is "large" is simply to note that the difference between the two mean values is large relative to the standard deviation.

There is nothing magical about the three thresholds of "small," "medium," or "large." Contrary to the Department's suggestion, these are not thresholds defined for any substantive purpose. Rather, they are simply arbitrary conventions. Cohen himself acknowledged the "danger of using terms like 'small', 'medium', and 'large' out of context." As further explained by Professor David Lane of Rice University:

> It is natural to ask what constitutes a large effect. Although there is no objective answer to this question, the guidelines suggested by Cohen (1988) stating that an effect size of 0.2 is a small effect, an effect size of 0.5 is a medium effect, and an effect size of 0.8 is a large effect have been widely adopted. Based on these guidelines, the effect size of 0.87 is a large effect.
>
> ***It should be noted, however, that these guidelines are somewhat arbitrary and have not been universally accepted.*** For example, Length (2001) argued that other important factors are ignored if Cohen's definition of effect size is used to choose a sample size to achieve a given level of power.[10]

[10] David Lane et al, Chapter 19 "Effect Size", Section 2 "Difference Between Two Means," (emphasis

There is thus no reason to consider 0.8 as "large" other than as an arbitrary convention that is sometimes but not always used.

Another way to see the arbitrariness of this characterization of a difference as "large" is to realize a Cohen's d test statistic of 0.8 means the difference is actually smaller than the standard deviation to which the difference is being compared. The ratio would have to be larger than 1.0 to mean the difference was larger than the standard deviation. Put another way, the cut off at 0.8 is a lower threshold than the old "one-standard deviation" (pattern) test used by Commerce, which itself has been criticized as too loose a standard. The cut off at 0.8 means that for a normal distribution, 44 percent of the transactions would fall outside the cut-off.[11] Intuitively, it is hard to see how something that happens almost half of the time is in any way unusual or "targeted." Such a low threshold is simply a way to find targeting where none truly exists.

Although these conventions are arbitrary, the real purpose of such a statistic is to standardize the size of the difference. As Professor Lane also explains:

> When the scale of a dependent variable is not inherently meaningful, it is common to consider the difference between means in standardized units. That is, effect size is measured in terms of the number of standard deviations the means differ by." [12]

---

added), http://onlinestatbook.com/2/effect_size/two_means.html. David Lane led the project team that developed this on-line text book, and is an associate professor of statistics at Rice University.

[11] David Lane et al, Chapter 19 "Effect Size", Section 2 "Difference Between Two Means," (emphasis added), http://onlinestatbook.com/2/effect_size/two_means.html.

[12] *See* David Lane et al, Chapter 19 "Effect Size", Section 2 "Difference Between Two Means," (emphasis added), http://onlinestatbook.com/2/effect_size/two_means.html.

This standardization is useful to provide some context to a variable for which the scale itself is not meaningful. Moreover, such standardization is critical when trying to compare the size of an effect across different studies. Thus effect sizes (as measured by Cohen's d test) "provide a standard metric for comparing across studies and thus are critical to meta-analysis {summarizing multiple studies on the same topic}." Indeed, that is one of the primary purposes of the Cohen's d test, a statistic that is "widely used in meta-analysis."[13]

Although it is certainly reasonable for the Department to want to have such a standardized measure of the degree of difference between mean values of its test group and comparison group, the Department needs to recognize the inherent limitations of this particular statistic. The convention on "large" is arbitrary. Moreover, the statistic does not in any way even try to address the more important task of distinguishing the true difference between the means and the statistical "noise" inherent in any set of data that varies. To say that the difference is "large" does not mean that the difference is "significant" in any meaningful sense of that word. The Cohen's d test might measure a difference that is greater than the convention of 0.8 as "large," but that measured difference might be completely unreliable and completely a construct of the small sample size and random noise in the data.

The Department had made its willingness to consider completely unreliable conclusions explicit in several recent decisions. The Department explicitly states it can use the Cohen's d test whenever it has "at least two observations" in its test and

[13] *See* http://en.wikiversity.org/wiki/Effect_size; *See also* http://en.wikiversity.org/wiki/Cohen%27s_d.

comparison groups. It is true that one can calculate a Cohen's d test when there are at least two observations in each group. But that is just another way of saying that whenever you have two numbers, you can calculate the mean and the standard deviation of those two numbers. Having at least two observations in each of two groups does not mean one can say anything meaningful – statistically meaningful, or meaningful in any other sense – about the difference between the means of those two groups.

The Cohen's d test is therefore not a substitute for the more traditional and widely recognized t-test for determining whether the difference between two mean values is truly a statistically significant difference. The Cohen's d test is not part of every standard introduction to statistical analysis because it is not a test of statistical significance. The Cohen's d test serves a narrower and more specialized purpose to standardize a measure of the size of a difference. In contrast, the t-test is taught in every introductory statistics course, because the t-test is in fact the basic way to measure whether the difference between two means is in fact a meaningful difference – a difference that is larger than zero and can be distinguished from the underlying noise in the data.

Since the t-test and the Cohen's d test address different issues, the Department should report both, and find the existence of differential pricing only when <u>both</u> standards have been met. Applying the t-test would allow the Department to determine whether the difference being observed is real, and not just a feature of the random variation in the data itself. Checking the Cohen's d test would then allow the Department to confirm that the difference is large enough to be considered evidence of possible targeted dumping. Indeed, this point has been stressed in discussions of how the Cohen's d test serves as a

compliment for, but not a replacement for, traditional tests of statistical significance. Reporting either one without the other is to ignore important information. The point is that a measured difference should be both statistically significant (the Department should be confident the difference is real and not an illusion of the data variability) and economically meaningful (the Department should be confident the difference is large in some meaningful way relative to the data).

Thus, the Department has identified a useful tool that might compliment other analysis using the traditional t-test to measure whether the difference between the two mean values is real or not. But the Department has not found a proper substitute for conducting a t-test to determine whether the differences being observed are in fact real and not artificial constructs of the random variations in the data.

Suggested Change *#2*: **Require two standard deviations from mean price**

The current differential pricing test applies too low a standard for a determining a significant price difference. Rather than the 0.8 standard deviation threshold currently being used, the Department should use a higher threshold. For the reasons discussed below, we believe a threshold of two standard deviations makes more sense in light of the statutory objective and sound policy.

At the outset, we note the 0.8 standard deviation threshold is arbitrary. Notwithstanding the warnings from Cohen himself about using 0.8 as a standard for determining "large" effect size[14], the Department has used this standard and has done so

[14] Cohen, *Statistical Power Analysis for the Behavioral Sciences*, ed II 1988.

without any explanation of why a different standard should be used in applying in the new differential pricing test than in applying prior targeted dumping tests, or why 0.8 is "large" in the context of targeted dumping other than the fact that it is often used in other contexts. Like with the issue of a one-tailed or two-tailed analysis, the use of 0.8 as a threshold instead of some other figure must have a rational basis related to what is being examined and for what purpose. Although a 0.8 threshold may be appropriate in certain circumstances (the Department has never adequately explained why it is reasonable in the differential pricing context), it may be equally inappropriate in other circumstances. We would note in this regard that in previous tests for targeted dumping the Department has used one standard deviation test, which would correspond to a 1.0 threshold when applying the Cohen's d test; meanwhile the traditional t-test approach to significance is based on approximately two standard deviations (1.96).

In statistics the standard deviation is used to measure the variation or dispersion from the average or mean. A low standard deviation indicates that the data points tend to be clustered near to the mean and a higher standard deviation indicates the data points are dispersed over a larger range. Using the normal distribution as the Cohen's d test assumes, one would expect 68.2% of the data points to be within one standard deviation (plus or minus) of the mean. Using this same normal distribution, one would expect 57.6% of the data points to be within plus or minus 0.8 standard deviations.[15] Thus, in a normal distribution the prices would be expected to vary from the standard deviation with 42.4% being outside 0.8 standard deviations and 31.8% being outside of one standard

[15] All calculations of means, standard deviations, and distributions were done using Excel instructions.

deviation. If the standard deviation is being used to test for whether prices differ significantly, the question is at what point relative to the standard deviation do prices differ significantly?

Because the Cohen's d procedure is comparing two sets of data, and the relative spread of two sets of data, the difference in the spread of the data is slightly different than when one is examining only a single set of data. The Cohen's d test result using 0.8 as the threshold would show 47.4% of the data points fall outside the 0.8 band - slightly less than half. At 1.0 it would show 55.4% outside the band - slightly more than half. The question is whether or not the prices inside and outside the 0.8 or 1.0 bands can as a group be considered to be at prices that are significantly different from each other when they are almost equal in terms of frequency of occurrence and, therefore, likelihood. In particular, keeping in mind that under the statute the Department must find a "pattern of prices" that differ significantly, in the case of what is nearly a 50-50 distribution it is difficult to argue that there is a pattern of any kind, let alone a pattern of prices that differ significantly. Indeed, if the prices on both sides of the 0.8 and/or 1.0 are clustered close to the standard deviation, there would be no pattern of prices that differ significantly. There might be some random variations, but a random variation is not part of a pattern. The amount of non-overlap needs to be higher to reasonably conclude there is truly a pattern of prices that differ significantly.

The table below illustrates the percent of non-overlap when applying various effect sizes using the Cohen's d test[16]:

Cohen's Standard	Effect Size	Percent of Non-overlap
	2.0	81.1%
	1.9	79.4%
	1.8	77.4%
	1.7	75.4%
	1.6	73.1%
	1.5	70.7%
	1.4	68.1%
	1.3	65.3%
	1.2	62.2%
	1.1	58.9%
	1.0	55.4%
	0.9	51.6%
Large	0.8	47.4%
	0.7	43.0%
	0.6	38.2%
Medium	0.5	33.0%
	0.4	27.4%
	0.3	21.3%
Small	0.2	14.7%
	0.1	7.7%
	0.0	0.0%

This table illustrates the importance of setting a higher threshold. Using either 0.8 (the current differential pricing test) or 1.0 (the old Nails test for targeted dumping) as the threshold is inappropriate. For either of those thresholds, the fact that sales are almost equally likely to be inside as outside the threshold makes it difficult to defend the threshold as one of true significance. This is not measuring anything extreme in the distribution that would render the sales outside either 0.8 or 1.0 significantly different. The appropriate threshold is the point at that the prices vary from the norm by a large or

[16] Effect Size (ES) – Effect Size Calculator (Lee Becker), University of Colorado Springs at Part II (www.uccs.edu/becker/effect-size).

significant amount. We do not see this until at least an effect size of 1.3 or 1.4[17] and it cannot be said to be a pattern until the effect size is at least 1.7.[18]

Overall, the traditional two standard deviation rule of thumb from a t-test provides the most reasonable threshold, with substantial enough non-overlap to provide a clear justification to conclude there is a pattern of prices that truly differ significantly. But even if the Department believes this threshold is too high, the threshold must be larger than 0.8 standard deviations.

To further illustrate the need for a much higher threshold, consider this example. We start with two datasets, with the means and standard deviations as listed below.[19]

Data	Test	Base
OBS1	116.0	111.0
OBS2	116.0	114.0
OBS3	118.0	113.0
OBS4	118.0	114.0
OBS5		113.0
OBS6		114.0
OBS7		113.0
OBS8		113.0
OBS9		128.0
Mean	117.0	114.78
Standard Deviation	1.0	4.76

In dataset 1 (the test set in this example), all observations are clustered around the mean but are 1.0 standard deviation from the mean. One could hardly characterize this as a

[17] With non-overlap at between 1.3 and 1.4 the Department would also be using a threshold consistent with its determination that a 66% ratio is sufficiently large to apply the alternative methodology to all sales.

[18] At 1.7 the variation reaches 75%.

[19] All calculations of means and standard deviations were done using Excel instructions.

pattern of pricing that differs significantly given that all observations despite being 1 standard deviation from the mean are within less than 2% of each other. Nor could one conclude that there is hidden dumping given the symmetry of the pricing pattern in data set 1 and the fact that all prices are equidistant from the mean. Yet, when compared with data set 2 (the base set) this test set still passes the 0.8 threshold for "large" used by the Department.

The above discussion simply demonstrates that use of 0.8 standard deviations can distort results when it is being used as a benchmark to determine whether differences are "large." We now turn to an example of the application of the Cohen's d test. When combined, the two datasets provided below would make up all prices in one CONNUM; the test set represents sales in the CONNUM to a particular customer, region or during a period of time.

Data	Test	Base
OBS1	116	118
OBS2	116	117
OBS3	118	118
OBS4	118	117
OBS5		118
OBS6		118
OBS7		118
OBS8		
OBS9		
OBS10		
OBS11		
Mean	117.0	117.71
Standard Deviation	1.0	0.45
Cohen's d		0.871

This example illustrates the defects in the Cohen's d test with a low 0.8 threshold. In this example, the test group with prices that have a 50% overlap with the base group (OBS3 and OBS4) and with a lowest priced observation is only 1 unit of measurement below the lowest price in the base group and less than 1% lower than that lowest price in the CONNUM, is nevertheless deemed to represent a pattern of prices which differ significantly by the Department when it applies Cohen's d using 0.8 as the standard of "large." In common sense terms the conclusion that the pattern of pricing in the test case differs significantly from the base is absurd notwithstanding that the difference between the two is considered "large" if the Department adopts the arbitrary standard of 0.8. Cohen himself has stressed that his definitions of "small", "medium", and "large" must be used with caution. This example illustrates why such caution is necessary and why the use of 0.8 has the potential of finding "large" differences when there are not large differences in any common sense meaning of that term. This absurd result, in turn, leads one to question whether or not use of Cohen's d with a 0.8 threshold is reasonable at all, as the Department claims.

Although one can likely also construct an example where Cohen's d might reveal a meaningful difference in prices sufficient to be "significant" in common sense terms, what is obvious from the above example is that the Cohen's d test applied mechanically using the 0.8 threshold cannot be relied upon to reveal a meaningful difference sufficient to be "significant." This, in turn, leads to a number of possible conclusions: (1) that the threshold for "large" should be significantly higher than 0.8; (2) that Cohen's d with a higher threshold should be used in combination with the t-test to ensure that the

Department is, in fact, making a determination of significance which is meaningful in a real sense; or (3) that Cohen's d is not a proper procedure to determine differential pricing.

In the event that the Department continues to use Cohen's d either by itself or in combination with the t-test, the Department should abandon the 0.8 threshold and move to a higher threshold. Indeed, LG submits that the **Department should adopt a test of 1.96 standard deviations** in its analysis. Adopting a 1.96 standard deviations approach would ensure that the Department is satisfying the statutory term "significantly." Adopting a test of 1.96 standard deviations would ensure compliance with the statutory standard.

Suggested Change *#3*: **Focus only on lower prices, not higher prices**

Although the Cohen's d test could be part of a reasonable approach for analyzing possible injurious differential pricing, this statistic still needs to be applied properly. Instead, the Department forgets the very context of the analysis and improperly considers the absolute value of the difference, instead of considering only positive values of the Cohen's d test. In other words, the Department's methodology allows higher priced U.S. sales transactions to the alleged target to provide evidence suggesting possible injurious differential pricing through lower priced U.S. sales. This approach makes absolutely no sense, given that the entire purpose of the AD law is to address lower priced U.S. sales.

This error is quite explicit in the SAS code. The DOC SAS code uses the absolute value of "d" – thereby over-stating the quantity of sales allegedly targeted. In other

words, both positive and negative value of the Cohen's d test are treated as equivalent or, stated differently, higher prices are being treated in the Department's margin program in the exact same manner as low prices: as targeted sales.

Although as a matter of mathematics a positive or negative number could be considered "large," the mathematics must be grounded in some context of what is being measured. The context here is checking for evidence of possible "differential pricing" that would mask dumping – in other words, looking for evidence that U.S. prices to one customer, region, or time period are lower than might otherwise be expected and are being masked by the use of an overall average. It makes no sense to look at U.S. prices that are higher than the comparison group, and use those higher prices – no matter how much higher – as evidence that harmful (within an antidumping context) differential pricing might be taking place. A higher price cannot possibly be evidence of harmful differential pricing because that means the differential pricing is higher than the base – precisely the opposite of the concern to be addressed by alleged targeted dumping.

Indeed, the absurdity of this approach is even more extreme. Under the current SAS code, the Department would consider slightly higher U.S. prices as not providing evidence of a large difference. If the mean U.S. price is somewhat higher, and the Cohen's d test calculates at 0.6 – in other words, the mean of the U.S. prices being tested is lower by an amount equal to 60% of the standard deviation – the Department would find no evidence of a "large" difference. But if the mean U.S. prices to the alleged target are much higher – 120% of the standard deviation – all of sudden now there is evidence of targeted dumping. The Cohen's d test would be -1.2 – minus because U.S. prices are

higher – and could turn into 1.2 after taking the absolute value. The 1.2 would be larger than 0.8, and the Department's methodology would consider that evidence of a large difference.

As documented in the Department's SAS code utilized in various recent AD investigations and AD reviews, the Department isn't comparing these higher prices to the whole universe of the CONNUM. Rather, the prices that the Department's own "test" classifies as aberrationally high are included in the comparison group for the sales dubbed aberrationally low. How can this be a permissible test? When the test is being applied to both the high and low price sales, how can both of those types of sales be included in the test group applied to the other?

Equally wrong in our view is the fact the Department's current methodology would find sufficient "differential pricing" to justify use of the alternative comparison methodology even if there are no prices below the mean that pass the Cohen's d test at 0.8, but there are prices above the mean that pass the Cohen's d test at 0.8. Based on the Department's definition of sales that constitute hidden or masked dumping (the prices of those sales passing Cohen's d below the 0.8 threshold), there are no sales which fall into the "hidden" or "masked" dumped category. In this instance, what is the purpose of putting this particular CONNUM in the "hidden" or "masked" dumping category when the Department's test does not find any sales on which hidden dumping was possible by virtue of using the average-to-average comparison? The purpose cannot be to unmask hidden dumping because there are no such sales in this situation where dumping was masked by use of the average-to-average comparison.

The Department has to date not provided any reasoned explanation either of why the definition of pattern of export prices used in the previous tests has been abandoned (i.e., why it has changed from a one tailed test focused on lower prices to a two tailed test that includes both lower prices and higher prices) or how these higher prices now being included in identifying the existence of targeted dumping or differential pricing are relevant to the existence of hidden dumping. We acknowledge that the statute does not explicitly require that the Department consider only lower priced sales in the differential pricing analysis. But neither does the statute require that the Department use both higher and lower priced sales.

Under the statute, however, the Department must adopt a test that is reasonable in light of the purpose of the provision when filling the gaps left by the statutory language and that provides the most accurate result. Although we agree that both high priced and low priced sales can contribute to a pattern of prices that differ significantly, the question remains whether the higher priced sales (i.e. those that pass Cohen's d at 0.8 or above) should be included in the universe of sales used to measure the existence of differential pricing (i.e. the 33% and 66% tests). The fact that higher priced sales implicitly through the calculation of the weighted average price can affect the outcome is already accounted for in using the mean of the weighted average sales prices in the Cohen's d procedure. Thus, the effect of these higher priced sales on masking dumping is already reflected in the use of the mean that includes these sales in the Cohen's d procedure. One can only have hidden dumping when there are sales below the Cohen's d threshold regardless of whether there are any sales above the Cohen's d threshold. This renders the inclusion of

sales above the Cohen's d threshold a meaningless measure of whether there is targeted dumping or differential pricing.

Therefore, the Department should adjust its methodology and consider only positive differences as possibly indicating a large difference that warrants further evaluation. This would simply be accomplished by deleting the "ABS" function in the SAS MARCO or Margin source code. Any differences which demonstrate that the test group has higher prices are inconsistent with the very premise of the alleged targeted dumping, and should therefore not be considered as evidence of possible targeting.

Indeed, we submit that not doing so would make the Department's interpretation contrary to the statute, for the reasons noted above.

Suggested Change *#4*: **Consider the three bases separately, not on an aggregated basis**

The statute permits the use of the alternative average-to-transaction methodology in situations in which it is demonstrated that prices differ significantly "among purchasers, regions, or time periods." In conducting this analysis, although it *separately* determines whether prices differ significantly by purchaser, by region, or by time period, the Department then aggregates the results of its application of the Cohen's d test for all three into a single amount and then uses this amount (i.e., the total or aggregate of all three bases allowing for use of the alternative comparison methodology) to determine whether the thresholds the Department has established for application of the alternative methodology are met. This methodology allows the Department to include in its ratio test, for example, sales to customers which pass the Cohen's d test, even if such sales are

trivial, simply because when added to sales by region and time period which pass Cohen's d they are included in the single thresholds (33% and 66%) used to determine whether differential pricing or targeted dumping is occurring.

LGE submits that this approach is wrong. The differential pricing exercise being undertaken by the Department is intended to "unmask" hidden dumping. Ironically, in doing so the Department is *masking* the fact that sales are not differentially priced by any of the individual bases identified in the statute. This is because when the Department uses an aggregate measure and does not examine customer, region, and time period separately, the results can differ. In essence, this approach captures sales that are not actually differentially priced by ignoring the distinctions made in the statute between three distinct situations: (1) differentially priced sales by customer; (2) differentially priced sales by region; and (3) differentially priced sales by time period.

This makes little sense. The 33%-66% thresholds should be applied to each individual basis. If 90% of the customers are not differentially priced, it makes no sense to add that 10% to some other basis to reach the 33% threshold. The 33%-66% thresholds implicitly assume they are from a universe of 100%, not the 300% that using an aggregated approach creates. Put another way, from a universe of 300%, the 33% threshold is saying that even when as few as 11% of the transactions are differentially priced, Commerce will ignore the basic rule of averaging. When only 11% of the transactions are priced differently, one cannot reasonably conclude that the 11% constitutes a pattern.

And therefore, Commerce should change it methodology and consider the three bases separately. The 33% and 66% threshold makes sense, but only when applied to each basis separately.

Suggested Change #5: **Do not exclude the test sales from the base sales**

Commerce currently compares the test group against other transactions excluding that test group. The logic for doing so, however, does not make sense in the context of testing for possible targeted dumping. The rule simply makes a finding of differential pricing more likely, and is thus an inappropriate bias to the testing.

The most frequent use of the Cohen's d test is to compare various types of treatments, as well as various levels of treatment, and the effect of those treatments on the same group (before and after) or two different groups (that is, one control group and one experimental group). For example, the Cohen's d test would be used to answer the question: "what is the relative size of the effect of taking aspirin on a regular basis with respect to the incidence of heart attacks?" In this classic two-group case, the comparison would be between one group taking the aspirin (experimental or test group) and another group taking a placebo (the control group). Another example would be what is the effect size of eliminating starches from a person's diet on that person's weight. In this case, the same group would be compared before eliminating starches and after eliminating starches to determine the effect size of eliminating starches. These cases involve external treatments and the same test group before and after the treatment or different groups with one subject to the treatment and the other controlled.

In the case of differential pricing, however, the context is very different. The Department is not attempting to measure the effect size of an external treatment but rather how the behavior of one group (the sales to a specific customer or the sales to a specific region or the sales during a specific time period) differs from the norm. Specifically, it is attempting to determine whether there is a pattern of pricing in a subset of sales that differs from the overall pattern of pricing. Although it is clear what constitutes the test group (sales to a particular customer or region or during a particular period), it is less clear what constitutes the base group against which the test group should be measured, called the "control" group above.

The Department has chosen, without any detailed explanation or justification of why, the base group as the universe of all sales (i.e., all sales in a particular CONNUM) minus those sales contained within the individual test group being tested. For example, in determining whether there is differential pricing in a particular period, the Department compares sales in one period (i.e. a specific quarter) with sales in all of the other periods (i.e., those quarters other than the one being tested). The question is: should the base group include or exclude the period being tested when being compared? This same question applies equally when testing by customer and region.

The purpose of determining whether there are patterns of prices that differ significantly is to determine whether or not the variation in price to certain groups differs from the normal pattern of prices so significantly as to allow normal comparisons to mask deviations in the test group pattern of pricing from "normal" pattern of pricing. The normal pattern of pricing by definition must include all the prices that occurred during the

period of investigation – including those for which a variation is being tested. Those transactions in the test group also occurred, and are also "normal" sales. The only legal basis for abandoning the average-to-average comparison is if there is a pattern of pricing that differs significantly from the pattern of pricing used in the average-to-average comparison. By excluding from the base group the sales to the test group, the Department is not comparing the pattern of pricing in the normal average-to-average comparison with the test group, but is comparing one test group (the mean of the prices of all sales excluding the prices from the test region, time period or customer) with another test group (the prices to a particular customer, region, or during a specific time period). The Department is not measuring variation of the sales in a particular test group from the mean of all sales (of a particular CONNUM, for a particular basis), but is instead measuring the differences in the variation of sales from one test group (the mean of all sales in a CONNUM minus the sales to a particular customer, region or period) to another test group (the particular customer, region, or period of the test group). The determination of whether prices differ significantly is only being measured as to two sub groups and not against the totality of prices used in the average-to-average comparison.

As an example, let's assume that an exporter has two customers. One accounts for 90% of the sales. The Department's methodology would apply the Cohen's d test by comparing those 90% of the sales with the other 10% of the sales and whether the mean for the 90% differs significantly from the other 10% of the sales. What if the 90% of the sales do pass the Cohen's d test and the difference is "large"? The 90% of the sales is the predominant factor in what the average prices are that are used in the average-to-

average comparison. Why is the fact that these sales "pass" the Cohen's d test in any way an indicator that there are significant differences in the prices that are somehow masking dumping? Since the Cohen's d test will also find a "large" difference when the 10% is measured, all of the sales in this example would pass the Cohen's d test and be deemed to represent differential pricing. By virtue of excluding the sales of the test customer from the base the Department ends up finding all sales in this particular CONNUM to be differentially priced. In short, using a base group that excludes the test group automatically skews the results and thus creates a bias. Such an approach is not reasonable.

An additional problem of excluding the test group from the base group is that the base group means being used in applying the Cohen's d test changes the bar for what is considered "normal" in each analysis. In the above example, the base group for determining the mean for one customer (e.g., the mean of the 90% of sales) is entirely different than the base group for determining the mean for the second customer (e.g., the mean of the 10% of sales). Similarly, in testing for whether or not prices differ significantly by time period, each time period is measured against a different base period (i.e., quarter 1 is measured against the means of quarters 2, 3, and 4, quarter 2 is measured against quarters 1, 3 and 4, etc.). The same is true for comparisons across regions.

Let's assume that we are attempting to determine whether or not a given basketball team (the Miami Heat) is taller on average than the average NBA team and whether that difference is large. The height of the average NBA team must include those

players that are part of the Miami Heat or the measurement being tested is not whether the Heat are taller than the average NBA team but whether the Heat are taller than all other NBA teams except the Heat. The differential pricing test should be testing whether prices to a given customer, region or during a particular time period are significantly different from the average price and, therefore, hide dumping. The average price necessarily must include the test group.

The average price should also be used because it is the average price, not the average price minus the test group, which is supposed to be distorting the results of the dumping comparison. The test conducted must relate to its objective. The objective is to determine whether the average price should or should not be used in the dumping comparison. The alternative is to use individual prices. As such, the question being answered by application of the Cohen's d test is whether there are significant differences in prices by region, customer, or time periods that are hidden by use of the average price of the CONNUM as a whole. The fact that the mean of two sub-groups of one CONNUM (i.e., one customer versus all other customers) differ significantly does not tell us whether this difference is hiding dumping because the mean of each of the two subgroups is not what is potentially hiding dumping. What is potentially hiding dumping is using the mean of the CONNUM. If there is not a "large" difference in the mean of each subgroup compared with the mean of the entire CONNUM, then the use of average-to-average is not masking anything because each subgroup mean is close to the mean being used in the average-to-average dumping comparison. This is true whether or not the difference of the means of the two subgroups is "large."

As a general concept, it is more reasonable to measure a subset of a larger group against its corresponding population – the entire population including the subset. In particular, when conducting multiple tests (in the case of measuring differential pricing each customer, time period and region are being tested) each should be measured against the same population (i.e., all sales in a given CONNUM for a given basis) such that the results of the test do not depend on the variability of the base group - but are measured more reliably against the norm set by the overall population being studied. Less variability in the base group results in a more accurate result across the various tests. For example, the extent to which prices to each of the customers buying a product designated under a given CONNUM differ significantly should not be determined based on a variable base group mean for each customer, but should be determined against the same base group mean for each customer. There is no justification either in the targeted dumping provision of the law or in statistics for using a variable mean when testing for differences.

In sum, Commerce should change its methodology and include the test sales in the universe of the comparison group. Based on the above, we believe that the correct calculation of the Cohen's d test statistic would be:

$$\frac{\bar{x}_{CONNUM-basis} - \bar{x}_{test}}{\sigma_{pooled}} - d$$

Where:

$\bar{x}_{test}$ Is defined are the mean net price of the test group for some basis and CONNUM.

$\bar{x}_{CONNUM-basis}$ Is defined as the mean net price of the CONNUM-basis combination.

σ_{pooled} Is defined as the standard deviation of the CONNUM-BASIS group only.

Suggested Change #6: **The variance should be based on a weighted average, not a simple average**

Beyond the other misuses of the Cohen's d test described above, the Department makes another basic mistake by calculating a simple average of the variance rather than a weighted average of the variance.

The SAS code makes clear that Commerce is using a simple average. The Department has two samples – the target and the non-target (or what the Commerce SAS code calls the "base group"). The SAS code makes clear that the Department calculates the pooled standard deviation as:

$$\sigma_{DOC} = \sqrt{\frac{\sigma_{base}^2 + \sigma_{target}^2}{2}}$$

But this formula represents a simple average, which treats the variance from the base group and the target group as equal, even when the two variances might be of very different sizes. Indeed, the Department's decision memoranda make clear it will apply this test as long as the comparison group "accounts for at least five percent of the total sales quantity of the comparable merchandise."[20] In other words, the Department contemplates using variances from samples of very different sizes.

The more appropriate formula for a weighted average that should be used in such a situation is the following:

$$\sigma_{correct} = \sqrt{\frac{(N_{base} - 1)\sigma_{base}^2 + (N_{target} - 1)\sigma_{target}^2}{N_{base} + N_{target} - 2}}$$

[20] *See e.g.*, *Xanthan Gum* note 9 supra, Post-Preliminary Analysis Memorandum at 3.

Where N_i is the total quantity sold to each group (i=target or base). The weighted average recognizes that the variance from two different samples of different sizes should have a different impact on the overall average variance. It should be noted, for the sake of consistency, that the weight average standard deviation actually simplifies to the straight-averaged one when N-Base=N-Target.

Consider for example that a respondent sells widgets to one "targeted" company. That company has 100 units purchased at an average price of $290. The non-targeted group has 900 transactions at an average price of $295. Is this $5 price difference large? Suppose further the standard deviation of the target group is $1 and for the non-target is $7.

Company	Volume	Avg. Price	Std. Dev.
Target	100	$290	$1
Non-target (base)	900	$295	$7

Using the Department's method the pooled standard deviation is calculated as:

$$\sigma_{DOC} = \sqrt{\frac{7^2 + 1^2}{2}} = \sqrt{50/2} = 5$$

However, the correct pooled standard deviation is:

$$\sigma_{correct} = \sqrt{\frac{(900-1)7^2 + (100-1)1^2}{900 + 100 - 2}} = \sqrt{\frac{(44051) + (99)}{998}} = 6.65$$

Using the correct pooled standard deviation, we can see in this example that the Department's methodology has unreasonably exaggerated the relative size of the difference:

Method	Difference in Means	Std. Dev.	Cohen's d
DOC current method	5	5	1
Correct method	5	6.65	0.75

In this example, the Department's "over weighing" of the target group (treating the 100 quantity as the same as the 900 quantity of the base) causes the Cohen's d test to exceed the Department's cutoff of 0.8. In fact, the correct Cohen's d test statistic is only 0.75 (which is less than the current Department cutoff of 0.8).

The use of a simple average is thus distorting, and gives too much weight to the variance from the target groups, which are often going to be smaller and with lower variance. In any event, whichever group is larger than the other group, the correct approach is a weighted average that properly accounts for differences in the sizes of the groups being compared.

Suggested Change #7: **Adjust the 25% change in margin test for low margins**

Commerce currently uses a 25% change in the margin as the sole test to determine if averaging cannot take into account price differences. Although this test may be reasonable for larger margins, that reasonableness breaks down for lower margins. It makes no sense to say that a 0.19 percent point change in a 0.75 percentage point margin

matters in this context, just because 0.19 percentage points is more than 25% of 0.75 percentage points. The 25% rule for the change in the margin should have a floor.

Using the same rules as currently in effect for *de minimis* margins, Commerce should ignore any change of the margin that would be considered *de minimis* and conclude that the averaging can take into account those price differences. This modification would mean that changes in the margin of less than 2.0 percentage points in an investigation or less than 0.5 percent in an administrative review would be deemed so small that the use of averaging can take into account those price differences. If the change in the margin is so small that it would be deemed *de minimis* on its own, then such a small change should be treated as not requiring departure from the rule to use average prices.

And therefore, LGE submits that Commerce should change its methodology and should ignore any change in the AD margin that would be less than the *de minimis* threshold when deciding whether averaging cannot take into account price differences.

Suggested Change #8: **Ensure sufficient number of transactions**

The current Commerce practice is to calculate a Cohen's d test any time there are more than four transactions in a CONNUM (two in the test group and two for which to test against). But such a small number of transactions simply do not permit any meaningful analysis of means, standard deviations, and differences. When there are only a handful of transactions, the means and standard deviations are too subject to distortion by outliers. The issue should not be having enough transactions for the mathematical

calculation of means and standard deviations to be possible. Rather, there should be enough transactions for those calculations to be both possible, but also meaningful.

A common rule of thumb for quantitative analysis is to require at least 30 data points. This rule of thumb makes sense given the use of the Cohen's d test, since that test assumes a normal distribution and that assumption requires a data set of certain size. This rule of thumb is also a common breakpoint for when statisticians begin to use special rules for small sample sizes. Assuming Commerce does not want to develop special rules to be used with smaller data sets, a more practical solution is to require at least 30 transactions in a CONNUM to apply the differential pricing tests.

Suggested Change #9: **Articulate clear definitions of "region," "time period" and "customer" in the context of the differential pricing analysis**

Commerce's definition of what constitutes a "region," "time period," or "customer" currently differs from one investigation to the next. While the statute grants Commerce the discretion to define these terms, Commerce should articulate clear definitions for each term so that interested parties have at least a basic level of predictability as to how Commerce will analyze pricing. The purpose of an antidumping duty order is to instill discipline in pricing decisions. Yet, respondents are not able to make informed pricing decisions under the order if they do not have clear, consistent guidance on the definitions that are so critical to Commerce s differential pricing analysis.

Under Commerce's current practice, "region" can be interpreted to hold a wide range of definitions. When comparing prices across multiple "regions," Commerce could define the "region" by geographical region (east, Midwest, west, south) or by state or by zip code or by any other geographic indication. This remarkably broad term is completely undefined. Thus, differential pricing could be found to exist under one definition of region, yet it could be absent using a second definition. A respondent has no ability to make informed pricing decisions where there is such variability in how those pricing decision will be analyzed by Commerce.

Respondents face similar unpredictability with respect to Commerce's analysis of pricing by "time period." Here, Commerce could analyze pricing by days, weeks, months, calendar quarters, or seemingly any other measure of time. Respondents are again faced with an exceptionally broad term and no guidance on how Commerce will apply that term from one case to the next.

LGE submits that Commerce should articulate clear definitions for these terms. Doing so would provide respondents with a measure of predictability for how pricing decision will be analyzed under the differential pricing analysis.

CONCLUSION

LGE respectfully urges the Department to take these comments into account when continuing to modify the particular methodologies used to implement 19 U.S.C. 1677f(d)(1)(B).